

Measuring star formation in high- z massive galaxies: A mid-infrared to submillimeter study of the GOODS NICMOS Survey sample

M. P. Viero^{1,2*}, L. Monceli^{1,3}, E. Mentuch^{4,2}, F. Buitrago^{5,6}, A. E. Bauer^{7,5},
E. L. Chapin⁸, C. J. Conselice⁵, M. J. Devlin⁹, M. Halpern⁸, G. Marsden⁸,
C. B. Netterfield^{2,10}, E. Pascale³, P. G. Pérez-González^{11,12}, M. Rex¹²,
D. Scott⁸, M. W. L. Smith³, M. D. P. Truch⁹, I. Trujillo^{13,14}, D. V. Wiebe⁸

¹ California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

² Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada

³ Department of Physics & Astronomy, Cardiff University, 5 The Parade, Cardiff, CF24 3AA, UK

⁴ Department of Physics & Astronomy, McMaster University, Hamilton, ON L8S 4M1

⁵ School of Physics and Astronomy, University of Nottingham, NG1 3AL, UK

⁶ SUPA†, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK

⁷ Australian Astronomical Observatory, PO Box 296, Epping, NSW 1710, Australia

⁸ Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada

⁹ Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104

¹⁰ Department of Physics, University of Toronto, 60 St. George Street, Toronto, ON M5S 1A7

¹¹ Departamento de Astrofísica, Facultad de CC. Físicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain

¹² Steward Observatory, The University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA

¹³ Instituto de Astrofísica de Canarias, E-38205, La Laguna, Tenerife, Spain

¹⁴ Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain

Accepted . Received 2011 April 18; in original form

ABSTRACT

We present measurements of the mean mid-infrared-to-submillimeter flux densities of massive ($M_\star \gtrsim 10^{11} M_\odot$) galaxies at redshifts $1.7 < z < 2.9$, obtained by stacking positions of known objects taken from the GOODS NICMOS Survey (GNS) catalog on maps at: 24 μm (*Spitzer*/MIPS); 70, 100, and 160 μm (*Herschel*/PACS); 250, 350, and 500 μm (BLAST); and 870 μm (LABOCA). A modified blackbody spectrum fit to the stacked flux densities indicates a median [interquartile] star-formation rate of $\text{SFR} = 63 [48, 81] M_\odot \text{yr}^{-1}$. We note that not properly accounting for correlations between bands when fitting stacked data can significantly bias the result. The galaxies are divided into two groups, disk-like and spheroid-like, according to their Sérsic indices, n . We find evidence that most of the star formation is occurring in $n \leq 2$ (disk-like) galaxies, with median [interquartile] $\text{SFR} = 122 [100, 150] M_\odot \text{yr}^{-1}$, while there are indications that the $n > 2$ (spheroid-like) population may be forming stars at a median [interquartile] $\text{SFR} = 14 [9, 20] M_\odot \text{yr}^{-1}$, if at all. Finally, we show that star formation is a plausible mechanism for size evolution in this population as a whole, but find only marginal evidence that it is what drives the expansion of the spheroid-like galaxies.

Key words: galaxies: evolution — galaxies: high-redshift — infrared: galaxies.

1 INTRODUCTION

The observed structural properties of massive galaxies ($M_\star \gtrsim 10^{11} M_\odot$) at high redshift ($z \gtrsim 1$) are difficult to reconcile with those of galaxies that populate the lo-

cal Universe. Most strikingly, they are on average much more compact in size than local galaxies of similar mass (Daddi et al. 2005; Trujillo et al. 2006). For the spheroid-like galaxy population, this size evolution has been particularly dramatic (a factor of 4–5 since $z \sim 2$, see e.g., Trujillo et al. 2007; Buitrago et al. 2008; Damjanov et al. 2009), with subsequent observations confirming these findings (e.g., Muzzin et al. 2009; Trujillo et al. 2011). Only

* E-mail: marco.viero@caltech.edu

† Scottish Universities Physics Alliance

a tiny fraction of massive galaxies in the local Universe have sizes comparable to those found at high redshift (Trujillo et al. 2009). The absence of similar mass counterparts in the local Universe (Trujillo et al. 2009) implies that some mechanism is acting on those high-redshift galaxies to make them grow in size (Hopkins et al. 2009; Bezanson et al. 2009).

In order to understand the mechanism responsible for this galaxy growth, a crucial point that needs to be addressed is the level of star formation (or star-formation rate: SFR) in this population. From an observational point of view, evidence for star formation in massive galaxies at high redshift is unclear, especially for the spheroid-like population. For example, small samples of high-quality spectroscopy (Kriek et al. 2006, 2009) find little or no star formation in this population; whereas, about 50% of these galaxies appear to have $24\mu\text{m}$ counterparts (Pérez-González et al. 2008), indicating an elevated level of star formation. This discrepancy may be due to biases inherent to their respective SFR estimators, which are either susceptible to errors in extinction correction and require deep spectroscopic observations; or probe emission from polycyclic aromatic hydrocarbons (PAHs), and thus provide a poor constraint on the thermal spectral energy distribution (SED).

An alternative probe of star formation is to observe in the far-infrared/submillimeter bands (FIR/submm), where emission is primarily from heated dust. It is known that in the local Universe the dust luminosity in star-forming regions is correlated with SFR (e.g., Kennicutt 1998; Chary & Elbaz 2001; Buat et al. 2007), with the most actively star-forming galaxies often the most dust obscured or even optically thick in the optical/UV (Genzel et al. 1998). Therefore, it is reasonable to expect that if high-redshift, compact, massive galaxies are vigorously forming stars, then they should be observable in the rest-frame FIR/submm.

However, due to the large beams of current submm telescopes, source confusion and flux boosting present significant obstacles to studying the star formation properties of anything other than the most luminous galaxies at high redshift (Moncelsi et al. 2011). For example, the $1\text{-}\sigma$ noise floor due to confusion in the $250\mu\text{m}$ band of *Herschel*/SPIRE is 5.8 mJy (Nguyen et al. 2010), which corresponds to the flux from galaxies at $z \sim 2$ with bolometric FIR luminosities of $L_{\text{FIR}} \sim 2 \times 10^{12} L_{\odot}$, i.e., ultra-luminous infrared galaxies (ULIRGs). As a result, a catalog of galaxies at $z > 2$ robustly detected above the confusion noise ($5\text{-}\sigma$) in the submm can only probe the bright end of the luminosity distribution. Stacking provides a mechanism to examine the full distribution, provided a reliable external catalog extending to faint fluxes is available (Marsden et al. 2009; Pascale et al. 2009).

In this work we perform a stacking analysis using a catalog of distant massive galaxies from the GOODS NICMOS Survey (GNS; Conselice et al. 2011) — which we select to have stellar masses $M_{\star} \geq 10^{11} M_{\odot}$ and redshifts $1.7 < z < 2.9$ — on maps from: *Spitzer*/MIPS (Rieke et al. 2004) at $24\mu\text{m}$; *Herschel*/PACS (Poglitsch et al. 2010) at 70, 100, and $160\mu\text{m}$; the Balloon-borne Large Aperture Submillimeter Telescope (BLAST; Devlin et al. 2009) at 250, 350, and $500\mu\text{m}$; and the Large APEX Bolometer Camera (LABOCA; Weiß et al. 2009) at $870\mu\text{m}$. Our objective is to estimate the average SFRs of high-redshift massive galaxies, and to look for differences between the disk-like and

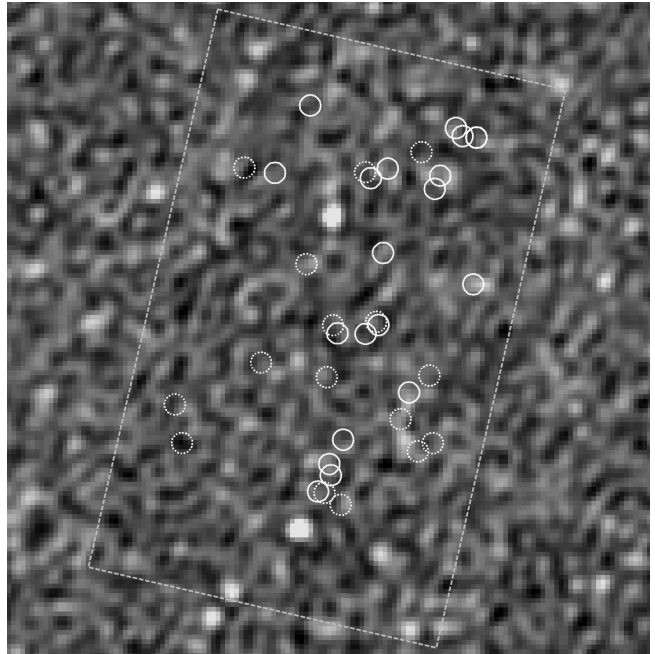


Figure 1. GNS catalog positions (white circles, $36''$ in diameter, solid are $n \leq 2$; dotted are $n > 2$) overlaid on a $20' \times 20'$ region of the BLAST $250\mu\text{m}$ map in GOODS-South. The overlapping *Herschel*/PACS region is outlined as a dashed box. The map has been convolved with a matched-filter (see Chapin et al. 2011) to help enhance the regions of submm emission. Most of the sources in our catalog lie along regions of faint emission. Note that the BLAST beam is many ($\sim 18\text{--}30$) times larger than a resolved galaxy, necessitating the stack. Furthermore, since the angular resolution of *Herschel*/SPIRE images will only improve by a factor of two, stacking will still be required to understand the FIR/submm properties of the faint population.

spheroid-like galaxies. An alternative approach, based on counterpart identification of similar GNS catalog sources, is carried out by Cava et al. (2010).

When required, we adopt the concordance model, a flat ΛCDM cosmology with $\Omega_{\text{M}} = 0.274$, $\Omega_{\Lambda} = 0.726$, $H_0 = 70.5\text{ km s}^{-1}\text{ Mpc}^{-1}$, and $\sigma_8 = 0.81$ (Hinshaw et al. 2009).

2 DATA

We perform our analysis on the Great Observatories Origins Deep Survey South field (GOODS-South), also known as the Extended Chandra Deep Field South (E-CDFS), which has field center coordinates $3^{\text{h}}32^{\text{m}}30^{\text{s}}, -27^{\circ}48'20''$. Here we briefly describe the catalog and maps.

2.1 Mass-Selected Catalog

Our catalog is the Buitrago et al. (2008) subset of the publicly available GOODS NICMOS Survey¹ (Conselice et al. 2011). Here we summarize its main features; for a more detailed description see Buitrago et al. (2008), Bluck et al. (2009) and Conselice et al. (2011). For details concerning

¹ <http://www.nottingham.ac.uk/astronomy/gns/index.html>

the data reduction procedure see Magee et al. (2007). The GNS is a large *HST* NICMOS-3 camera program of 60 H -band pointings (180 orbits), with limiting magnitudes of $H \sim 26.8$ (5σ), optimized to collect data for as many massive ($M_\star \gtrsim 10^{11} M_\odot$) galaxies as possible at high redshift ($1.7 < z < 2.9$), making it the largest sample of such galaxies to date. Of these, 36 are in the southern field for which we have infrared and submm maps.

Redshifts and stellar masses of these objects are calculated using the BVRiizJHK filters. Photometric redshifts are found using standard techniques (e.g., Conselice et al. 2007), while spectroscopic redshifts for 7 objects are compiled from the literature (Wuyts et al. 2008; Popesso et al. 2009; Balestra et al. 2010). Stellar masses of these objects are estimated by fitting the multi-color photometry to model SEDs — produced with stellar population synthesis models — resulting in uncertainties of ~ 0.2 dex (e.g., Bundy et al. 2006).

Additionally, due to the excellent depth and resolution of the NICMOS images (pixel scale after resampling of $0.''1 \text{ pixel}^{-1}$, and a point spread function [PSF] of $0.''3$ full width half maximum [FWHM]), we are able to estimate the Sérsic (1968) indices and sizes of the objects using the GALFIT code (Peng et al. 2002). Average properties of the sources used in our analysis are listed in Table 1.

The selection for the GNS galaxies is based on mass and redshift, with $1.7 < z < 2.8$. These galaxies were located initially through color selection techniques, such as the BzK (Daddi et al. 2007), ERO (Yan et al. 2004) and DRG Papovich et al. (2006) criteria, and later refined through spectroscopic and photometric redshifts within the two GOODS fields. Conselice et al. (2011) perform several tests to ensure that the sample is complete. A possible bias might be that extremely dusty galaxies could be missed by this criteria due to attenuation, but the deep limiting H -band magnitude greatly exceeds that of the expected upper bound for dusty SMGs (~ 23.3 mag, Frayer et al. 2004), so that we are confident that we are not missing the dustiest galaxies due to attenuation. Lastly, it is expected that this selection of galaxies closely approximates the true ratio of red to blue galaxies in these mass and redshift ranges. For more details concerning the selection technique and possible biases (see Conselice et al. 2011).

2.2 Spitzer

We use the publicly available *Spitzer*/MIPS map at $24 \mu\text{m}$ from the Far Infrared Deep Extragalactic Legacy Survey (FIDEL²). The 5σ point source sensitivity of this map is 0.03 mJy .

2.3 PACS

We use publicly available *Herschel*/PACS (Poglitsch et al. 2010) observations of the GOODS-South field from the PACS Evolutionary Probe³ (PEP; Lutz et al. 2011) survey. The data was re-processed with the *Herschel* Processing Environment (HIPE; Ott 2010, continuous integration build

number 6.0.2110). PEP was designed to provide data in all three PACS bands. Since PACS can only observe in two bands simultaneously — at $160 \mu\text{m}$ (red) and either $70 \mu\text{m}$ (blue) or $100 \mu\text{m}$ (green) — we use two sets of observations to produce maps at all three wavelengths. We combine the available deep observations using the standard PACS pipeline, choosing a high-pass filter parameter of 20 for the blue and green bands, and 30 for the red band (corresponding to suppression of scales larger than 40 and $60''$ on the sky, respectively; see Müller et al. 2011). In order to prevent ringing effects around bright sources caused by the high-pass filter the pipeline performs an initial crude reduction and automatically masks out the brightest sources in the subsequent iterations of de-glitching and filtering. The r.m.s. depths of the final maps are 0.31 , 0.44 , and 1.5 mJy at 70 , 100 , and $160 \mu\text{m}$, respectively.

As reported by Müller et al. (2011), the relatively strong high-pass filter adopted along with the masking of the bright sources may attenuate the final photometry of faint sources. To account for these effects, we produce maps of a few, isolated, unmasked, faint point sources of different flux density, using the same parameters as were used in the reduction of the GOODS-South maps; we then mask these sources and create new maps. We use the average ratio of the flux densities of the same sources in the two maps as our estimate of the attenuation factor due to the high-pass filter. We find that the magnitude of the attenuation mildly increases for increasing wavelengths, as expected given the shape of the $1/f$ noise over the relevant frequency range ($\propto f^{-0.5}$; Lutz et al. 2011). The estimated attenuation factors are 0.80 , 0.78 , and 0.75 at 70 , 100 , and $160 \mu\text{m}$, respectively. Note that a slightly different approach was followed by Lutz et al. (2011), who perform tests on the red band by adding simulated sources to the timelines before masking and high-pass filtering; they find that the filtering modifies the fluxes by 16% for very faint unmasked point sources. Despite the slight disagreement with our finding at $160 \mu\text{m}$, and because of the lack of an estimate for the blue and green bands from the PEP team, we choose to adopt our three estimated factors for consistency.

2.4 BLAST

The BLAST maps in GOODS-South⁴ consist of a deep region covering $\sim 0.9 \text{ deg}^2$ which completely encompasses the southern sources in the Buitrago et al. (2008) catalog (Figure 1), and have r.m.s. depths of 11 , 9 , and 6 mJy , at 250 , 350 , and $500 \mu\text{m}$, respectively (Devlin et al. 2009). Due to large instrumental beams (36 , 42 , and $60''$) and steep source counts (approximately following $dN/dS \propto S^{-3}$; Patanchon et al. 2009), source confusion contributes substantially to the noise in these maps, and is estimated to be $\sigma_{\text{confusion}} \approx 21$, 17 , and 15 mJy in the three bands (Marsden et al. 2009). The BLAST maps were made using a naive mapmaker (Pascale et al. 2011). Further details on the instrument may be found in Pascale et al. (2008), while flight performance and calibration are provided in Truch et al. (2009).

² http://data.spitzer.caltech.edu/popular/fidel/20070917_enhanced/docs/fidel_dr2.html/

³ <http://www.mpe.mpg.de/ir/Research/PEP/>

⁴ Available at: <http://blastexperiment.info/results.php>

	N	z_{median}	z_{iqr}	M_{\star} (M_{\odot})	R_e (kpc)	n	T (K)	L_{FIR} (L_{\odot})	SFR ($M_{\odot} \text{ yr}^{-1}$)
All	36	2.285	1.980–2.500	1.85×10^{11}	2.00	2.03	$29.4^{+1.4}_{-0.8}$ [27.3, 31.6]	$6.2^{+1.1}_{-1.0}$ $[4.7, 8.0] \times 10^{11}$	63^{+11}_{-11} [48, 81]
$n \leq 2$	20	2.285	2.085–2.500	1.93×10^{11}	2.43	1.05	$32.6^{+1.0}_{-0.4}$ [30.8, 34.6]	$12.0^{+1.4}_{-1.5}$ $[9.8, 14.8] \times 10^{11}$	122^{+15}_{-15} [100, 150]
$n > 2$	16	2.270	1.865–2.625	1.74×10^{11}	1.49	3.25	$27.6^{+0.3}_{-7.6}$ [24.2, 30.8]	$1.4^{+0.2}_{-0.8}$ $[0.9, 2.0] \times 10^{11}$	14^{+2}_{-8} [9, 20]

Table 1. Average properties of stacked samples. R_e is the effective radius. Dust temperatures, bolometric FIR luminosities and SFRs, corrected to a Chabrier (2003) IMF, are shown with the corresponding upper and lower Gaussian uncertainties, and interquartile ranges in square brackets (see Section 4 for details).

2.5 LABOCA

The LABOCA E-CDFS Submm Survey (LESS; Weiß et al. 2009) provides deep 870 μm data, with an r.m.s. depth to better than 1.2 mJy across the full $30' \times 30'$ field, with an effective resolution of $27''$ FWHM. For a detailed description of the instrument see Siringo et al. (2009).

3 METHOD

3.1 Stacking Formalism

Stacking is a well established technique for finding the average properties of objects which individually are undetectable by using external knowledge of their positions in a map (e.g., Dole et al. 2006; Wang et al. 2006; Marsden et al. 2009; Pascale et al. 2009). We follow the formalism of Marsden et al. (2009, hereafter M09), to which we refer to for a full description of the stacking method. Here we summarize the salient features of the technique.

M09 showed that the mean flux density of an external catalog is simply the covariance of the mean-subtracted map with the catalog, divided by the variance of the catalog density. If the catalog is Poisson-distributed, then a powerful diagnostic is that the variance of the source density should equal the mean, and the average flux density can be re-written as the mean map value at the position of each catalog source. This is true no matter what the size of the beam or surface density of sources in the map, so long as the sources are uncorrelated at the scale of the beam. The algorithm has been extensively tested with Monte Carlo simulations on mock random maps with increasing source densities, and was shown consistently to recover the correct mean flux density, with no dependence on the number of sources per beam (Figure 2). If however the catalog is clustered on the beam scale, the stacked flux will be biased high, compared to the properly normalized covariance, by a factor equal to the catalog variance at the beam scale divided by the mean source density. In the following section we show that this factor is consistent with unity for our data.

Uncertainties and possible biases of our measurement are estimated by generating random catalogs and stacking them on the actual maps themselves. We find that the uncertainties are Gaussian-distributed and scale as the map r.m.s. (including confusion noise) divided by the square root of the number of catalog entries. Note that these uncertainties account for both instrumental and source confusion noise, as well as for any pixel-pixel correlations introduced by the map-making algorithm (e.g., the “drizzling” technique used to produce the PACS maps with the standard pipeline).

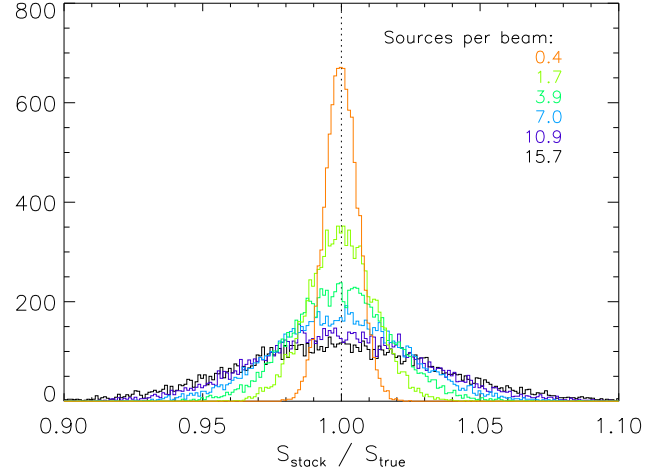


Figure 2. Histograms showing the ratio of recovered stacked fluxes to true flux for 10,000 simulations. The stacks were performed on simulated 0.25 deg^2 maps based on a random catalog of 12,500 sources, with size and source densities typical for deep $24 \mu\text{m}$ MIPS catalogs. We have repeated the test for six beam sizes in the range $10\text{--}60''$, which probe the effects of stacking at source densities ranging from 0.4 to 16 sources per beam. As described in Section 3.1 and in M09, larger beams lead to larger uncertainties, but in all cases, the stacked values are consistent with the true catalog flux, showing that there is no bias when stacking on uncorrelated catalogs.

3.2 Testing the Poisson Hypothesis

Stacking provides an unbiased estimate of the mean flux only when the sources in the sky are uncorrelated. While massive galaxies have been shown to cluster quite strongly (e.g., Foucaud et al. 2010), we find that on scales relevant for this analysis they are essentially Poisson-distributed, as we show with the following tests:

1) In the presence of clustering, the FWHM of the postage-stamp of stacked sources would be larger than the nominal instrumental PSF. We compare our measured stacked $24 \mu\text{m}$ PSF to that measured from stacking the sources used in M09 (Magnelli et al. 2009), which were shown to be Poisson distributed (see Figure 3 of M09), and find that they are identical.

2) If the sources are Poisson-distributed over a given scale, then by definition the average number of sources in a cell of that size should equal the variance. We test that by dividing the field into equal sized cells, from 2.7 to $0.225'$ on a side, and find that the ratio of the variance to the mean is consistent with unity at all scales.

3) In the presence of strong clustering around massive galaxies we would expect to find more sources per beam surrounding the galaxies than would be found at random. We calculate the number of sources inside a BLAST beam radius at the locations of each massive galaxy and compare that to what we would expect at random. From 1,000 Monte Carlo simulations we find 1.10 ± 0.13 , 1.16 ± 0.17 , and 1.28 ± 0.21 sources per beam at 250, 350, and 500 μm , compared to the measured 1.04, 1.13, and 1.17, respectively. We extend this test to galaxies with $\log(M_*/M_\odot) > 9$, to account for the possibility of less massive galaxies clustering around our more massive ones. We find there are 2.85 ± 0.40 , 3.83 ± 0.51 , and 5.97 ± 0.73 sources per beam at 250, 350, and 500 μm , compared to the measured 2.53, 4.04, and 5.87, respectively. Thus, while there are multiple sources per beam at all wavelengths, because their distribution is consistent with Poisson, they do not bias the result.

There still remains the possibility, however, that even fainter ($< 13 \mu\text{Jy}$ at 24 μm), undetected sources cluster around detected ones. We can estimate their potential contribution in the following way. If clustered, faint sources contribute significantly to the stacked flux density for large beams then after convolving the 24 μm map (whose beam is $6''$) with a much larger beam we would expect the stacked flux density to increase. On the other hand, as described in the previous section, if the faint sources are Poisson-distributed then we would expect only the noise to increase. We find that after convolving the 24 μm map with a $60''$ beam, the stacked flux density per source is $0.08 \pm 0.11 \text{ mJy}$, compared to the original $0.081 \pm 0.005 \text{ mJy}$ (see Table 3). Thus, the stacked signal does not change, but the errors increase substantially, which is consistent with what we would expect from additional, Poisson-distributed sources in the beam. We therefore conclude that the contribution from faint clustered sources is negligible.

3.3 SED Fitting, IR Luminosities, and Star-Formation Rates

We model the thermal dust emission as a modified black body with an SED of the form;

$$S_\nu = A\nu^\beta B(\nu, T), \quad (1)$$

where $B(\nu, T)$ is the blackbody spectrum with amplitude A , and β is the emissivity index, which effectively takes into account the variability of dust temperatures within a single galaxy. Following Blain et al. (2003) we set β to 1.5. Additionally, we replace the mid-infrared exponential on the Wien side of the spectrum with a power-law of the form $f_\nu \propto \nu^{-\alpha}$ (with $\alpha = 2$, following Blain et al. 2003), which in practice means imposing that the two functions and their first derivatives be equal at the transition frequency. In our case that transition occurs at rest-frame $\sim 74 \mu\text{m}$ (for $T \sim 30 \text{ K}$; see § 4.3).

Our SED fitting procedure estimates the amplitude and temperature of the above template, keeping α and β fixed. For the BLAST points, the SED fitting procedure (described in detail in Chapin et al. 2008) takes the width and shape of the photometric bands into account, as well as the absolute photometric calibration uncertainty in each band (see Truch et al. 2009). Correlations due to instrumental noise

Band (μm)	24	70	100	160	250	350	500	870
24	1	0.11	0.13	0.23	0.35	0.28	0.22	0.05
70		1	0.92	0.77	0.22	0.15	0.08	0.006
100			1	0.86	0.27	0.19	0.11	0.007
160				1	0.44	0.33	0.20	0.04
250					1	0.70	0.62	0.11
350						1	0.70	0.14
500							1	0.13
870								1

Table 2. Pearson correlation matrix for all bands.

are estimated and accounted for with a Monte Carlo procedure. Because we do not possess similar detailed data for *Spitzer*/MIPS and LABOCA, these photometric points are not color-corrected, whereas we do apply a color-correction to the PACS points, following the standard procedure described in Müller et al. (2011, see their Table 4.2, for a power law ν^{-2}); the color-correction factors are 1.016, 1.012, 1.017 at 70, 100, and 160 μm , respectively, and have a negligible impact on the final results. The PACS points are assumed to have completely uncorrelated instrumental noise among bands.

Correlated confusion noise must also be accounted for in the fit, as these correlations reduce the significance of a combination of single band detections. We estimate the Pearson coefficients of the correlation matrix for all bands (see Table 2) from the beam-convolved maps within a region of 0.064 deg^2 that encompasses all the sources in the GOODS-South NICMOS catalog. In § 4.3 we will show how the effect of correlations between bands is quite significant, especially among PACS and BLAST bands (see also Monceli et al. 2011), and thus including them in the SED fitting algorithm is crucial.

SEDs are corrected for redshift by assuming the median redshift for each subset (see column 3, Table 1). Interquartile errors reflecting the uncertainty in dimming due to the width of the redshift bin are estimated with a Monte Carlo, where 1000 mock redshifts with the same distribution as the chosen subset (i.e., all, disk-like, and spheroid-like) are drawn, and the dimming factor for each redshift is calculated.

The resulting infrared luminosity, L_{FIR} , is conventionally the integral of the rest-frame SED between 8 and 1000 μm , and the SFR is estimated as

$$\text{SFR} [M_\odot \text{ yr}^{-1}] = 1.728 \times 10^{-10} \times L_{\text{FIR}} [L_\odot], \quad (2)$$

from Kennicutt (1998), which assumes a Salpeter (1955) initial mass function (IMF). To convert to a Chabrier (2003) IMF, $\log(\text{SFR})$ must be corrected by lowering 0.23 dex (e.g., Kriek et al. 2009; van Dokkum et al. 2010).

4 RESULTS

4.1 Stacking Results

Stacked flux densities and $1-\sigma$ uncertainties are reported in the second column of Table 3. We find statistically significant, non-zero signals in all the submm bands, with 2-, 3-, 3-, and 4- σ detections at 250, 350, 500, and 870 μm , respec-

Band (μm)	All (mJy/source)	$n \leq 2$ (disk-like) (mJy/source)	$n > 2$ (spheroid-like) (mJy/source)
24	0.081 ± 0.005	0.130 ± 0.007	0.020 ± 0.007
70	0.16 ± 0.07	0.36 ± 0.09	-0.05 ± 0.10
100	0.39 ± 0.09	0.84 ± 0.13	-0.17 ± 0.14
160	1.2 ± 0.3	2.9 ± 0.5	-0.66 ± 0.50
250	5.0 ± 2.9	9.3 ± 3.9	-0.3 ± 4.4
350	7.9 ± 2.3	10.7 ± 3.1	4.5 ± 3.5
500	5.3 ± 1.9	6.2 ± 2.6	4.2 ± 2.9
870	0.97 ± 0.26	1.03 ± 0.35	0.9 ± 0.4

Table 3. The mean flux densities of massive galaxies in the GNS catalog from stacking. Reported are the results for all of the sources, as well as those identified as disk-like and spheroid-like, based on their Sérsic indices, n .

tively, as well as robust 16-, 3-, 4-, 4- σ detections at 24, 70, 100, 160 μm , respectively.

Next, we divide the catalog by Sérsic index into: those with $n > 2$, which are spheroid-like and thus more likely to have suppressed star formation; and those with $n \leq 2$, which are disk-like and thus more likely to be actively forming stars (Ravindranath et al. 2004). Contamination of one population into the other has been shown with simulations to be very low ($< 10\%$; Buitrago et al. 2011), but when galaxies do cross the $n = 2$ threshold, it is always from $n < 2$ to $n > 2$, i.e., from spheroid-like to disk-like.

The results are listed in the third and fourth columns of Table 3. At 24 μm , we measure a distinct signal from both populations, with 19- σ and 3- σ detections from the disk-like and spheroid-like sources, respectively. At longer wavelengths, for the disk-like population we detect signals with greater significance than that of the combined catalog, between 2.5- σ and 6.5- σ , in each FIR/submm band; whereas for the spheroid-like population we find a much weaker signal, with four bands consistent with zero.

While the error on the stacks is Gaussian, the uncertainty associated with the average rest-frame L_{FIR} is dominated by the width of the redshift distribution, which is not Gaussian. Thus, for estimating T , L_{FIR} , and SFR (Section 4.3), we choose to adopt the median value and interquartile range, as they best reflect the asymmetric shape of the redshift distribution, which ultimately determines the uncertainty of our measurement. For reference we also quote the Gaussian uncertainties. We anticipate that the lower Gaussian errors on T , L_{FIR} , and SFR for the spheroid-like subset exceed the lower bound of the interquartile range, and reflect the elevated level of uncertainty in our measurement.

4.2 Contribution of stellar emission

At $z \sim 2.3$ the observed 24 μm band probes rest-frame wavelengths of 6–7 μm , which in addition to PAH emission, is where the Rayleigh-Jeans tail of stellar emission lies. It is then plausible that the emission we find in this band may be entirely attributed to stellar emission. Since our detection of PAH and dust emission, particularly for the spheroid population, is supported strongly by this data point (given its high signal-to-noise), in this section, we investigate the contribution of pure stellar emission to the observed 24 μm band.

Any additional emission not attributed to stellar emission is likely associated with PAH emission, which in turn is accompanied by longer wavelength dust emission that we have inferred our SFRs from. We note that this emission may be associated with either star forming regions or evolving main sequence stars (such as AGB and TP-AGB stars). Typically, emission associated with star formation dominates in most galaxies (even those with moderate SFRs) as the infrared light-to-mass ratio is up to three orders of magnitude larger for a simple stellar population (SSP) of 10^7 yrs compared to an SSP of 10^9 yrs (Piovan et al. 2006), where TP-AGB emission would be most significant.

To investigate the contribution of pure stellar emission in our sample, we calculate the predicted 24 μm observed flux densities from stellar population synthesis models using redshifts and stellar masses as per our catalog (see Section 2.1). We opt to use a galaxy template with solar metallicity and an exponentially declining SFR with an e-folding time of 500 Myr, generated with the stellar population synthesis code PEGASE.2 (Fioc & Rocca-Volmerange 1997). Output from non-stellar emission or evolving main-sequence stars is not included, as the source of non-stellar emission at 7 μm is assumed to be the same as that of the FIR emission. Assuming a formation redshift of $z = 9$, the galaxy ages range from 1.5 to 3 Gyr and the predicted 24 μm flux densities due to stellar emission range from 1.3 to 8.8 μJy , depending primarily on the galaxy’s redshift. For each stacked sample, we find the predicted contamination per galaxy from stellar emission is 3.0, 2.9, and 3.9 μJy for the entire sample, the disk-like and spheroid-like populations, respectively. Contributing at most 20% (see Table 3) to the observed 24 μm flux densities of the spheroid population and less than 5% to the observed 24 μm flux densities of the disk-like and total sample, we conclude that the mid-infrared observations (restframe 7–8 μm) included in our analysis are dominated by non-stellar emission (i.e., dust and PAH emission).

4.3 Best-fit SEDs and Star-Formation Rates

The best-fit SED and interquartile range to the stacked values of the complete catalog are shown in the left panel of Figure 3, corresponding to a median (plus/minus Gaussian) [interquartile] temperature of $T = 29.4^{+1.4}_{-0.8}$ [27.3, 31.6] K, luminosity of $L_{\text{FIR}} = 6.2^{+1.1}_{-1.0}$ [4.7, 8.0] $\times 10^{11} L_{\odot}$, and SFR = 63^{+11}_{-11} [48, 81] $\text{M}_{\odot} \text{yr}^{-1}$.

As a sanity check, we compare our modified blackbody approximation to the best-fit template of Chary & Elbaz (2001, hereafter CE01). The purpose of this is simply to reassure ourselves that an exponential approximation on the Wien side of the thermal SED is not an unreasonable way to estimate the contribution to the bolometric luminosity short of the SED peak, rather than an attempt to derive SFRs from fitting SED templates. Thus, for each of the 101 templates, we approximate the stacked SED by taking the average of templates shifted to the redshift of each galaxy in the catalog; this acts to smear out the otherwise highly-variable PAH region of the rest-frame SED probed by the 24 μm band. We fit the resulting template to our photometric points without accounting for calibration uncertainties, color corrections, or correlations among bands. The best-fit template is shown as a 3-dot-dashed line in Figure 3, and falls well inside our error region. However, the SFR of the

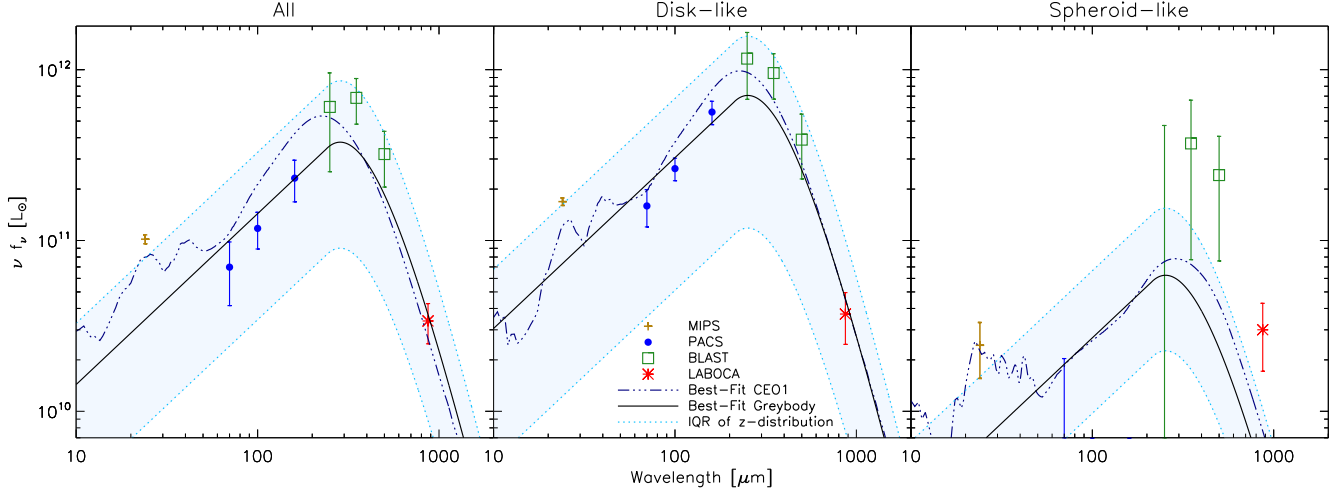


Figure 3. SED fits to the stacked flux densities of all (left), disk-like (center), and spheroid-like (right) sources. The median value of the redshift distribution, $z \sim 2.3$, is used here to convert flux densities into rest-frame luminosity. The brown crosses are from Spitzer ($24 \mu\text{m}$); the blue dots are from PACS (70 , 100 , and $160 \mu\text{m}$); the green squares are from BLAST (250 , 350 , and $500 \mu\text{m}$); and the red asterisks are from LABOCA ($870 \mu\text{m}$). The error bars represent the $1\text{-}\sigma$ Gaussian uncertainties from the stacks as listed in Table 3. The SED is modeled as a modified black body with a fixed emissivity index $\beta = 1.5$, and a power-law approximation on the Wien side with slope $\alpha = 2$. The solid black lines are the best-fit SEDs, while the dotted light-blue lines enclosing the shaded regions show the uncertainties due to the width of the redshift distribution (interquartile range), which clearly dominate over the Gaussian errors on the stacks (see Section 4.1). The navy 3-dot-dashed lines are the best-fit, redshift-averaged templates from Chary & Elbaz (2001).

best-fit template is $\text{SFR} = 87 \text{ M}_\odot \text{ yr}^{-1}$, or $\sim 38\%$ larger than our modified blackbody estimate. This overestimate likely arises because the fit with the CE01 template does not include the substantial correlations among bands (see Section 3.3), which reduce the significance of the combination of individual photometric points.

We then separately fit the stacked flux densities measured for disk-like and spheroid-like galaxies. The best-fit modified blackbody SED for the disk-like population is shown in the center panel of Figure 3, and results in a median (plus/minus Gaussian) [interquartile] temperature of $T = 32.6^{+1.0}_{-0.4}$ [30.8, 34.6] K, luminosity of $L_{\text{FIR}} = 12.0^{+1.4}_{-1.5}$ [9.8, 14.8] $\times 10^{11} L_\odot$, and $\text{SFR} = 122^{+15}_{-15}$ [100, 150] $\text{M}_\odot \text{ yr}^{-1}$. The best-fit CE01 template is also shown, and corresponds to a $\text{SFR} = 142 \text{ M}_\odot \text{ yr}^{-1}$.

Likewise, the best-fit modified blackbody SED for the spheroid-like population is shown in the right panel of Figure 3, and results in a median (plus/minus Gaussian) [interquartile] temperature of $T = 27.6^{+0.3}_{-7.6}$ [24.2, 30.8] K, luminosity of $L_{\text{FIR}} = 1.4^{+0.2}_{-0.8}$ [0.9, 2.0] $\times 10^{11} L_\odot$, and $\text{SFR} = 14^{+2}_{-8}$ [9, 20] $\text{M}_\odot \text{ yr}^{-1}$. Note that the lower Gaussian errors exceed the lower bound of the interquartile range, thus reflecting the elevated level of uncertainty in our measurement. Once again, the best-fit CE01 template is shown, which corresponds to a $\text{SFR} = 16 \text{ M}_\odot \text{ yr}^{-1}$.

Finally, to check that contributions to the rest-frame SED from PAHs, which are highly variable, are not significantly influencing the best-fit result, we re-fit the modified black body after excluding: i) just the $24 \mu\text{m}$ point; and ii) all points below rest-frame $100 \mu\text{m}$. In the first scenario we find $\text{SFR} = 57^{+9}_{-14}$ [42, 72], 109^{+11}_{-18} [91, 135], and 12^{+5}_{-7} [8, 17] $\text{M}_\odot \text{ yr}^{-1}$, while in the second scenario we find $\text{SFR} = 67^{+11}_{-16}$ [50, 83], 129^{+16}_{-23} [107, 160], and 30^{+10}_{-7} [19, 42] $\text{M}_\odot \text{ yr}^{-1}$, for all, disk-like, and spheroid-like galaxies, respectively. In the first case, the SFRs decrease

only marginally, and within the error bars, suggesting that the $24 \mu\text{m}$ point alone does not unreasonably influence the result. In the second case, SFRs for all and disk-like galaxies are mildly affected, while the spheroid-like galaxies are artificially boosted by a factor of two simply because we have removed the two data points consistent with zero.

Thus, although the best-fit SED to the combined stack returns a robust, $4\text{-}\sigma$ detection, it is clear that signal is dominated by the disk-like, $n \leq 2$ galaxies, which are detected at $5\text{-}\sigma$. The best-fit to the spheroid-like, $n > 2$ galaxies, on the other hand, returns a marginal $2\text{-}\sigma$ result, which suggests, but does not formally detect, a low level of star formation taking place in the spheroid-like population.

We note that if correlations between bands are not properly accounted for when finding the best-fit SED, the corresponding SFRs are 94, 163, and $32 \text{ M}_\odot \text{ yr}^{-1}$ for all, disk-like, and sphere-like galaxies, respectively. This is significantly different, and the reason is that if correlations between bands are not considered, more weight is attributed to the BLAST measurements than is appropriate, pulling the best-fit up. Intuitively this makes sense: since confusion noise arises from multiple sources in a beam, a larger beam has more sources in it and thus more variance, i.e., more confusion noise. Of course, for bands of similar wavelengths those sources are more or less present in each map, resulting in confusion noise that is not independent. Though this will be less of a problem for *Herschel*/SPIRE, the beam size and thus improvement in confusion noise is only of order $\sim 2\times$, so that correctly accounting for correlated confusion noise will still be very important.

5 DISCUSSION

5.1 Consequences for Galaxy Growth

There are indications that massive galaxies at high redshift are the cores of present-day massive ellipticals (Hopkins et al. 2009; Bezanson et al. 2009), and that the growth of these galaxies takes place mostly in the outskirts via star formation and minor mergers (Hopkins et al. 2009; van Dokkum et al. 2010) — a process sometimes referred to as “inside-out” growth, which has also been observed in hydrodynamical cosmological simulations (Naab et al. 2009; Johansson et al. 2009; Oser et al. 2010). Furthermore, van Dokkum et al. (2010) find that a SFR of $55 \pm 13 \text{ M}_{\odot} \text{ yr}^{-1}$ at $z \sim 2$ is necessary to account for the mass growth they observe in massive galaxies selected by number density, from $z = 2$ to the present day, and that for $z \gtrsim 1.5$ the mechanism for growth is primarily star formation. Note that nearly half of their $z \sim 2$ subsample of massive galaxies has $n < 2$ (see right panel of their Figure 7) — a fraction similar to our own. Our measurement of $63 [48, 81] \text{ M}_{\odot} \text{ yr}^{-1}$ for the entire sample agrees well with their finding; however, we do not find convincing evidence that star formation is the mechanism driving the expansion in spheroid-like galaxies.

5.2 Potential contribution from other sources of dust heating

Star formation may not be the only explanation for infrared emission in our sample which consists of very massive, yet relatively young systems. The age of the universe by $z = 3$ – 1.8 is just ~ 1.5 – 3 Gyr, providing a strict upper limit on the ages of the stellar populations. If these galaxies formed the bulk of their stellar mass, as their colors suggest, early on, then it is likely that they contain a large population of stars undergoing post-main-sequence phases in which carbonaceous dusty material is being produced and heated by very luminous stars. While it is generally accepted in the current versions of stellar population synthesis models (Maraston 2005; Bruzual 2010; Conroy & Gunn 2010) that thermally-pulsating asymptotic giant branch (TP-AGB) stars can contribute up to 70% of the emission seen in the near-infrared bands at ages of 1–2 Gyr, there has been little work calibrating the global contribution of this population to a galaxy’s infrared luminosity. By extension, given the masses and ages of our galaxies, we cannot rule out the possibility that the infrared emission we have detected in our analysis is partially due to dust heated and created by post-main-sequence stars.

5.3 Red and Dead?

Our best-fit SED to stacked data does not correspond to a formal detection of star formation in the spheroid-like ($n > 2$) galaxies, however, the high $24 \mu\text{m}$ flux might indicate a non-zero star-formation rate. Though we have stated that $24 \mu\text{m}$ emission alone is insufficient for accurately estimating the *level* of star formation in a galaxy, locally, $24 \mu\text{m}$ emission is typically well correlated with star-forming regions (Calzetti et al. 2007; Kennicutt et al. 2009). Additionally, emission from evolved stars seems unable to account for the level of $24 \mu\text{m}$ emission observed (§ 4.2). Therefore,

it seems plausible that star formation may be occurring in these galaxies at some level. Furthermore, if a low level of star formation does indeed exist, given the noise properties of our maps, the only bands which would permit a significant detection are the 24 and $870 \mu\text{m}$ bands — those in which our measurements have signal-to-noise greater than 2.5.

If star formation is occurring in the spheroid-like galaxies, even at a low level, and if they are fair analogs of the apparently red-and-dead compact spheroids seen by e.g., Kriek et al. (2009), then why is it that star formation is not significant in ultra-deep spectroscopy? One possibility is that the star formation is localized in very dust-obscured regions. Note that although Kriek et al. (2009) detect a faint $\text{H}\alpha$ line, concluding that SFRs are at most 2 – $4 \text{ M}_{\odot} \text{ yr}^{-1}$, that is after correcting for a very moderate amount of extinction ($A_v = 0$ – 0.3 mag). For this galaxy to actually be forming around $14 \text{ M}_{\odot} \text{ yr}^{-1}$, $L_{\text{H}\alpha}$ would need to have been underestimated by a factor of ~ 3.5 – 7 , which corresponds to 1.4 – 2.1 mag of extinction. Considering that resolved observations of nearby galaxies showing extinction values of $A_{\text{H}\alpha} > 3$ are common in H II regions (Prescott et al. 2007) and regions of high star formation (Mentuch et al. 2010), this amount of extinction is not unrealistic.

Lastly, note that our low levels of observed star formation are in disagreement with Cava et al. (2010), who (after correcting by 0.23 dex due to differences in the assumed IMF) find an average SFR of 30 – $60 \text{ M}_{\odot} \text{ yr}^{-1}$ for the spheroid-like galaxies. Their average SFRs are based on photometry of individual galaxies at $24 \mu\text{m}$, and at 250 , 350 , and $500 \mu\text{m}$ from *Herschel*/SPIRE with a mean detection fraction for the spheroid-like population of ~ 0.4 at $24 \mu\text{m}$ and ~ 0.15 at $250 \mu\text{m}$. This selection makes it difficult to properly compare measurements.

6 SUMMARY

Our goal was to search for evidence of star formation in high-redshift massive galaxies, with the hope of leading to a better understanding of the mechanisms responsible for their growth. We found that on average the full catalog of sources are forming stars with a median [interquartile] SFR = $63 [48, 81] \text{ M}_{\odot} \text{ yr}^{-1}$, which can be decomposed into a relatively strong signal for the disk-like galaxies, with a median [interquartile] SFR = $122 [100, 150] \text{ M}_{\odot} \text{ yr}^{-1}$, and a marginal signal for the spheroid-like population, with a median [interquartile] SFR = $14 [9, 20] \text{ M}_{\odot} \text{ yr}^{-1}$.

The level of star-formation detected for the full catalog is in good agreement with other measurements of galaxy growth (e.g., van Dokkum et al. 2010) which show that star formation can account for most of the growth at these redshifts. However, despite having detected stacked emission at 24 and $870 \mu\text{m}$, we are unable to say convincingly that star formation is responsible for the dramatic size evolution of the spheroid-like population.

Lastly, though a red sequence appears to already be in place by $z \sim 2$ (Kriek et al. 2009), we found hints that perhaps the red, compact, spheroid-like galaxies may not be completely dead. Future stacking work with larger catalogs and better maps will go a long way to further understanding this question. Better data bracketing the peak with SPIRE (250 , 350 , and $500 \mu\text{m}$; Griffin et al. 2010) will make more

robust estimates of the SED possible, and will greatly increase our understanding of star formation in high-redshift massive galaxies.

ACKNOWLEDGMENTS

The authors would like to extend a big thanks to the LESS team for providing the 870 μm map, and to David Frayer for the *Spitzer* 70 μm map used in the first version of this paper. We would also like to thank Ian Smail, Kimberly Scott, and Ivana Damjanov for constructive comments. Finally, we would like to thank Hervé Aussel for his helpful advice. We acknowledge the support of NASA through grant numbers NAG5-12785, NAG5-13301, and NNGO-6GI11G, the NSF Office of Polar Programs, the Canadian Space Agency, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the UK Science and Technology Facilities Council (STFC). This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX) under programme numbers 078.F-9028(A), 079.F-9500(A), 080.A-3023(A), and 081.F-9500(A). APEX is a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory.

REFERENCES

- Balestra I., Mainieri V., Popesso P., et al., 2010, *A&A*, 512, A12
- Bezanson R., van Dokkum P. G., Tal T., Marchesini D., Kriek M., Franx M., Coppi P., 2009, *ApJ*, 697, 1290
- Blain A. W., Barnard V. E., Chapman S. C., 2003, *MNRAS*, 338, 733
- Bluck A. F. L., Conselice C. J., Bouwens R. J., Daddi E., Dickinson M., Papovich C., Yan H., 2009, *MNRAS*, 394, L51
- Bruzual G., 2010, in G. Bruzual & S. Charlot ed., *IAU Symposium Vol. 262 of IAU Symposium, Population Synthesis: Challenges for the Next Decade*. pp 55–64
- Buat V., Takeuchi T. T., Iglesias-Páramo J., et al., 2007, *ApJS*, 173, 404
- Buitrago F., Trujillo I., Conselice C. J., Bouwens R. J., Dickinson M., Yan H., 2008, *ApJLetters*, 687, L61
- Buitrago F., Trujillo I., Conselice C. J., Haeussler B., 2011, *ArXiv* 1111.6993
- Bundy K., Ellis R. S., Conselice C. J., et al., 2006, *ApJ*, 651, 120
- Calzetti D., Kennicutt R. C., Engelbracht C. W., et al., 2007, *ApJ*, 666, 870
- Cava A., Rodighiero G., Pérez-Fournon I., et al., 2010, *MNRAS*, 409, L19
- Chabrier G., 2003, *PASP*, 115, 763
- Chapin E. L., Ade P. A. R., Bock J. J., et al., 2008, *ApJ*, 681, 428
- Chapin E. L., Chapman S. C., Coppin K. E., et al., 2011, *MNRAS*, 411, 505
- Chary R., Elbaz D., 2001, *ApJ*, 556, 562
- Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
- Conselice C. J., Bluck A. F. L., Buitrago F., et al., 2011, *MNRAS*, 413, 80
- Conselice C. J., Bundy K., Trujillo I., et al., 2007, *MNRAS*, 381, 962
- Daddi E., Dickinson M., Chary R., et al., 2005, *ApJLetters*, 631, L13
- Daddi E., Dickinson M., Morrison G., et al., 2007, *ApJ*, 670, 156
- Damjanov I., McCarthy P. J., Abraham R. G., et al., 2009, *ApJ*, 695, 101
- Devlin M. J., Ade P. A. R., Aretxaga I., et al., 2009, *Nature*, 458, 737
- Dole H., Lagache G., Puget J.-L., et al., 2006, *A&A*, 451, 417
- Fioc M., Rocca-Volmerange B., 1997, *A&A*, 326, 950
- Foucaud S., Conselice C. J., Hartley W. G., Lane K. P., Bamford S. P., Almaini O., Bundy K., 2010, *MNRAS*, 406, 147
- Frayer D. T., Reddy N. A., Armus L., Blain A. W., Scoville N. Z., Smail I., 2004, *AJ*, 127, 728
- Genzel R., Lutz D., Sturm E., et al., 1998, *ApJ*, 498, 579
- Griffin M. J., Abergel A., Abreu A., et al., 2010, *A&A*, 518, L3+
- Hinshaw G., Weiland J. L., Hill R. S., et al., 2009, *ApJS*, 180, 225
- Hopkins P. F., Bundy K., Murray N., Quataert E., Lauer T. R., Ma C., 2009, *MNRAS*, 398, 898
- Johansson P. H., Naab T., Ostriker J. P., 2009, *ApJLetters*, 697, L38
- Kennicutt Jr. R. C., 1998, *ARAA*, 36, 189
- Kennicutt Jr. R. C., Hao C.-N., Calzetti D., et al., 2009, *ApJ*, 703, 1672
- Kriek M., van Dokkum P. G., Franx M., Illingworth G. D., Magee D. K., 2009, *ApJLetters*, 705, L71
- Kriek M., van Dokkum P. G., Franx M., et al., 2006, *ApJLetters*, 649, L71
- Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, *ApJ*, 700, 221
- Lutz D., Poglitsch A., Altieri B., et al., 2011, *ArXiv* 1106.3285
- Magee D. K., Bouwens R. J., Illingworth G. D., 2007, in R. A. Shaw, F. Hill, & D. J. Bell ed., *Astronomical Data Analysis Software and Systems XVI Vol. 376 of Astronomical Society of the Pacific Conference Series, NICRED: A NICMOS Image Processing Pipeline*. p. 261
- Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2009, *A&A*, 496, 57
- Maraston C., 2005, *MNRAS*, 362, 799
- Marsden G., Ade P. A. R., Bock J. J., et al., 2009, *ApJ*, 707, 1729
- Mentuch E., Abraham R. G., Zibetti S., 2010, *ApJ*, 725, 1971
- Monceli L., Ade P. A. R., Chapin E. L., et al., 2011, *ApJ*, 727, 83
- Müller T., Nielbock M., Balog Z., Klaas U., Vilenius E., 2011, *Technical Report PACC-ME-TN-037 v1.0, PACS Photometer - Point-Source Flux Calibration*
- Muzzin A., van Dokkum P., Franx M., Marchesini D., Kriek M., Labbé I., 2009, *ApJLetters*, 706, L188
- Naab T., Johansson P. H., Ostriker J. P., 2009, *ApJLetters*, 699, L178
- Nguyen H. T., Schulz B., Levenson L., et al., 2010, *A&A*,

- 518, L5+
- Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, *ApJ*, 725, 2312
- Ott S., 2010, in *Astronomical Data Analysis Software and Systems XIX* Vol. 434 of *Astronomical Society of the Pacific Conference Series*, The Herschel Data Processing System - HIPE and Pipelines - Up and Running Since the Start of the Mission. pp 139–+
- Papovich C., Moustakas L. A., Dickinson M., et al., 2006, *ApJ*, 640, 92
- Pascale E., Ade P. A. R., Bock J. J., et al., 2008, *ApJ*, 681, 400
- Pascale E., Ade P. A. R., Bock J. J., et al., 2009, *ApJ*, 707, 1740
- Pascale E., Auld R., Dariush A., et al., 2011, *MNRAS*, 415, 911
- Patanchon G., Ade P. A. R., Bock J. J., et al., 2009, *ApJ*, 707, 1750
- Peng C. Y., Ho L. C., Impey C. D., Rix H., 2002, *AJ*, 124, 266
- Pérez-González P. G., Trujillo I., Barro G., Gallego J., Zamorano J., Conselice C. J., 2008, *ApJ*, 687, 50
- Piovan L., Tantaló R., Chiosi C., 2006, *MNRAS*, 366, 923
- Poglitsch A., Waelkens C., Geis N., et al., 2010, *A&A*, 518, L2+
- Popesso P., Dickinson M., Nonino M., et al., 2009, *A&A*, 494, 443
- Prescott M. K. M., Kennicutt Jr. R. C., Bendo G. J., et al., 2007, *ApJ*, 668, 182
- Ravindranath S., Ferguson H. C., Conselice C., et al., 2004, *ApJLetters*, 604, L9
- Rieke G. H., Young E. T., Engelbracht C. W., et al., 2004, *ApJS*, 154, 25
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sersic J. L., 1968, *Atlas de galaxias australes. Observatorio Astronómico*
- Siringo G., Kreysa E., Kovács A., et al., 2009, *A&A*, 497, 945
- Truch M. D. P., Ade P. A. R., Bock J. J., et al., 2009, *ApJ*, 707, 1723
- Trujillo I., Cenarro A. J., de Lorenzo-Cáceres A., Vazdekis A., de la Rosa I. G., Cava A., 2009, *ApJLetters*, 692, L118
- Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, *MNRAS*, 382, 109
- Trujillo I., Ferreras I., de la Rosa I. G., 2011, *ArXiv* 1102.3398
- Trujillo I., Feulner G., Goranova Y., et al., 2006, *MNRAS*, 373, L36
- van Dokkum P. G., Whitaker K. E., Brammer G., et al., 2010, *ApJ*, 709, 1018
- Wang W., Cowie L. L., Barger A. J., 2006, *ApJ*, 647, 74
- Wei A., Kovács A., Coppin K., et al., 2009, *ApJ*, 707, 1201
- Wuyts S., Labbé I., Schreiber N. M. F., Franx M., Rudnick G., Brammer G. B., van Dokkum P. G., 2008, *ApJ*, 682, 985
- Yan H., Dickinson M., Eisenhardt P. R. M., et al., 2004, *ApJ*, 616, 63